О ВЕРОЯТНОСТНЫХ СВОЙСТВАХ СТАТИСТИЧЕСКОЙ ОЦЕНКИ МНОГОМЕРНОЙ ЭНТРОПИИ ШЕННОНА

В. Ю. Палуха, Ю. С. Харин

НИИ прикладных проблем математики и информатики БГУ
Минск, Беларусь
E-mail: palukha@bsu.by, kharin@bsu.by

Исследованы вероятностные свойства приращения частотной статистической оценки *n*-мерной энтропии Шеннона при увеличении длины фрагмента. Получены явные формулы моментов 1-го и 2-го порядка приращения оценки энтропии в предположении, что наблюдаемая последовательность является чисто случайной. Также получена формула математического ожидания статистической оценки энтропии для наблюдаемой конечной реализации равномерно распределённой случайной последовательности.

Ключевые слова: многомерная энтропия, равномерно распределённая случайная последовательность, нормальное распределение, математическое ожидание.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

В криптографии и теории информации часто возникает задача статистического оценивания энтропии Шеннона [1] выходной последовательности криптографического генератора. Являясь мерой неопределённости, энтропия лежит в основе критериев качества генераторов случайных и псевдослучайных последовательностей. Для оценки качества генератора в смысле его близости по свойствам к равномерно распределённой случайной последовательности (РРСП) необходимо знание вероятностных свойств оценок энтропии РРСП. Дадим вначале определение *n*-мерной энтропии Шеннона.

Будем рассматривать двоичные последовательности $\{x_t\} \in V = \{0,1\}$ на некотором вероятностном пространстве (Ω, *F*, *P*). Пусть $p_{i_1,\dots,i_n} = P\{x_{t+1} = i_1, \dots, x_{t+n} = i_n\}$ – распределение вероятностей *n*-граммы $(x_{t+1}, \dots, x_{t+n}) \in V_n$, которое предполагается не зависящим от $t \in \boldsymbol{N}_0 = \boldsymbol{N} \cup \{0\}$. Многомерная (*n*-мерная) энтропия Шеннона для фрагмента длины *n* равна [1]:

$$H\{x_1, \dots, x_n\} = h(n) = -\sum_{i_1,\dots,i_n \in V} p_{i_1,\dots,i_n} \ln p_{i_1,\dots,i_n}. \qquad (1)$$

Обозначим: $i = \sum_{j=1}^{n} 2^{j-1} i_j$ – представление числа $i \in \{0, 1, \dots, 2^n - 1\}$ в двоичной системе счисления, $p_i(n) = P\{\sum_{j=1}^{n} 2^{j-1} x_j = i\} = p_{i_1,\dots,i_n}$, $i = 0, \dots, 2^n - 1$. Тогда формула (1) примет вид

$$h(n) = -\sum_{i=0}^{2^n-1} p_i(n) \ln p_i(n). \tag{2}$$

Пусть наблюдается последовательность $x_1, \ldots, x_T \in V$. Для удобства «зациклим» последовательность до длины $T + n - 1$: $x_{T+1} = x_1$, …, $x_{T+n-1} = x_{n-1}$. Обозначим $X^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)}) = (x_k, \ldots, x_{k+n-1}) \in V_n$, $k = \overline{1, T}$. Построим частотные оценки распределения вероятностей $\{p_i(n)\}$, $i = 0, \ldots, 2^n - 1$:

$$\hat{p}_i(n) = \frac{1}{T}\sum_{k=1}^{T} \delta_{\overline{X}^{(k)},i}, \quad \overline{X}^{(k)} = \sum_{j=1}^{n} 2^{j-1} x_j^{(k)}, \quad \delta_{\overline{X}^{(k)},i} = \begin{cases} 1, \overline{X}^{(k)} = i; \\ 0, \overline{X}^{(k)} \neq i. \end{cases} \tag{3}$$

Используя подстановочный принцип, построим статистическую оценку энтропии (2) с использованием оценок (3):

$$\hat{h}(T, n) = -\sum_{i=0}^{2^n-1} \hat{p}_i(n) \ln \hat{p}_i(n). \tag{4}$$

В [2] для проверки гипотезы о близости распределения вероятностей наблюдаемой последовательности $\{x_t\}$ к распределению вероятностей чисто случайной последовательности предложено использовать приращение оценки энтропии (4):

$$G(T, n) = \hat{h}(T, n) - \hat{h}(T, n-1), \quad G(T, 1) = \hat{h}(T, 1), \tag{5}$$

однако строгого обоснования предлагаемого критерия не представлено.

ПРИРАЩЕНИЕ ЭНТРОПИИ

Обозначим: $\lambda = \frac{T}{2^n}$; $\Pi_1(\lambda), \Pi_2(\lambda)$ – независимые одинаково распределённые по закону Пуассона с параметром $\lambda > 0$ случайные величины;

$$f(u_1, u_2) = -u_1 \ln u_1 - u_2 \ln u_2 + (u_1 + u_2) \ln(u_1 + u_2); \tag{6}$$

$$\mu(\lambda) = E\{\Pi_1(\lambda) \ln \Pi_1(\lambda)\}; \tag{7}$$

$$\gamma = \frac{1}{2\lambda} \operatorname{cov}\{f(\Pi_1(\lambda), \Pi_2(\lambda)), \Pi_1(\lambda) + \Pi_2(\lambda)\}; \tag{8}$$

$$\zeta(\lambda) = E\{\Pi_1(\lambda) \ln \Pi_1(\lambda)(\Pi_1(\lambda) - \lambda)\}; \tag{9}$$

$$g(u_1, u_2) = f(u_1, u_2) - E\{f(\Pi_1(\lambda), \Pi_2(\lambda))\} - \gamma(\Pi_1(\lambda) + \Pi_2(\lambda) - 2\lambda); \tag{10}$$

$$\sigma^2(\lambda) = D\{\Pi_1(\lambda) \ln \Pi_1(\lambda)\} = E\{\Pi_1^2(\lambda) \ln^2 \Pi_1(\lambda)\} - \mu^2(\lambda); \tag{11}$$

$$\nu(\lambda) = \operatorname{cov}\{(\Pi_1(\lambda) + \Pi_2(\lambda)) \ln(\Pi_1(\lambda) + \Pi_2(\lambda)), \Pi_1(\lambda) \ln \Pi_1(\lambda)\}. \tag{12}$$

Лемма. Для величин (7), (12) справедливы равенства:

$$\mu(\lambda) = e^{-\lambda} \lambda \sum_{k=1}^{+\infty} \frac{\lambda^k \ln(k+1)}{k!}; \tag{13}$$

$$\nu(\lambda) = e^{-2\lambda} \lambda \sum_{k=1}^{+\infty} \frac{\lambda^k \ln(k+1)}{k!} \sum_{j=1}^{k} (j+1) \ln(j+1) C_{k+1}^{j+1} - \mu(2\lambda)\mu(\lambda). \tag{14}$$

Доказательство. Докажем вначале (13), используя определение математического ожидания и распределение вероятностей Пуассона:

$$\mu(\lambda) = E\left\{\Pi_1(\lambda)\ln\Pi_1(\lambda)\right\} = \sum_{k=0}^{+\infty} k\ln k\,\frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}\sum_{k=2}^{+\infty}\frac{\lambda^k\ln k}{(k-1)!} = e^{-\lambda}\lambda\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}.$$

Теперь докажем (14). Проводя эквивалентные преобразования над (12) с учётом независимости случайных величин $\Pi_1(\lambda)$ и $\Pi_2(\lambda)$, получим

$$\begin{aligned}
&\nu(\lambda) = \operatorname{cov}\left\{\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\ln\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right), \Pi_1(\lambda)\ln\Pi_1(\lambda)\right\} = \\
&= E\left\{\left(\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\ln\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)-\mu(2\lambda)\right)\left(\Pi_1(\lambda)\ln\Pi_1(\lambda)-\mu(\lambda)\right)\right\} = \\
&= E\left\{\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\ln\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\Pi_1(\lambda)\ln\Pi_1(\lambda)\right\} - \mu(2\lambda)\mu(\lambda) - \mu(2\lambda)\mu(\lambda) + \\
&+\mu(2\lambda)\mu(\lambda) = E\left\{\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\ln\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\Pi_1(\lambda)\ln\Pi_1(\lambda)\right\} - \mu(2\lambda)\mu(\lambda).
\end{aligned} \tag{15}$$

Для распределения вероятностей Пуассона имеем

$$\begin{aligned}
&P\left\{\Pi_1(\lambda) = j, \Pi_1(\lambda)+\Pi_2(\lambda) = k\right\} = P\left\{\Pi_1(\lambda) = j, \Pi_2(\lambda) = k-j\right\} = \\
&= \frac{e^{-\lambda}\lambda^j}{j!}\cdot\frac{e^{-\lambda}\lambda^{k-j}}{(k-j)!} = \frac{e^{-2\lambda}\lambda^k}{j!(k-j)!} = \frac{k!}{j!(k-j)!}\cdot\frac{e^{-2\lambda}\lambda^k}{k!} = C_k^j\,\frac{e^{-2\lambda}\lambda^k}{k!}.
\end{aligned} \tag{16}$$

С учётом (16) и принятых обозначений получим

$$\begin{aligned}
&E\left\{\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\ln\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\Pi_1(\lambda)\ln\Pi_1(\lambda)\right\} = \sum_{k=0}^{+\infty}\sum_{j=0}^{k} kj\ln k\ln j\,\frac{e^{-2\lambda}\lambda^k}{k!}C_k^j = \\
&= e^{-2\lambda}\sum_{k=2}^{+\infty} k\ln k\,\frac{\lambda^k}{k!}\sum_{j=2}^{k} C_k^j\, j\ln j = e^{-2\lambda}\lambda\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}\sum_{j=1}^{k}(j+1)\ln(j+1)C_{k+1}^{j+1}.
\end{aligned} \tag{17}$$

Из (17) и (15) следует (14). ■

Распределение вероятностей статистики (5) описывается следующей теоремой [2].

Теорема 1. При истинной гипотезе $H_* = \{\{x_t\} - \text{РРСП}\}$ статистика (5) имеет асимптотически нормальное распределение $N_1(\mu_T, \sigma_*^2)$ при $T, n \to +\infty$ с математическим ожиданием μ_T и дисперсией σ_*^2, где

$$G_*(T,n) = \mu_T = \frac{1}{2\lambda}E\left\{f\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)\right\}, \tag{18}$$

$$\sigma_*^2 = \frac{\sigma_T^2}{T^2} = \frac{2^{n-1}}{T^2}D\left\{g\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)\right\}, \quad \sigma_T^2 = 2^{n-1}D\left\{g\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)\right\}. \tag{19}$$

В [2] параметры нормального распределения приведены лишь в общем виде. Однако для математического ожидания и дисперсии можно вывести явные формулы.

Теорема 2. Для величин (18) и (19) справедливы равенства:

$$\mu_T = e^{-\lambda}\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}\left(e^{-\lambda}2^k - 1\right), \tag{20}$$

$$\begin{aligned}
\sigma_*^2 = \frac{e^{-\lambda}}{T}\Bigg(&\sum_{k=1}^{+\infty}(k+1)\ln^2(k+1)\frac{\lambda^k}{k!}\left(2^k e^{-\lambda}+2\right) - 2\lambda e^{-\lambda}\left(\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}\left(2^k e^{-\lambda}-1\right)\right)^2 - \\
&- 2e^{-\lambda}\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}\sum_{j=1}^{k}(j+1)\ln(j+1)C_{k+1}^{j+1} -
\end{aligned} \tag{21}$$

$$-e^{-\lambda}\left(\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}\left(2^k e^{-\lambda}(k+1-2\lambda)-2(k+1-\lambda)\right)\right)^2.$$

Доказательство. Докажем вначале (20), используя (6) и (7):

$$E\left\{f\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)\right\}=$$

$$=E\left\{-\Pi_1(\lambda)\ln\Pi_1(\lambda)-\Pi_2(\lambda)\ln\Pi_2(\lambda)+(\Pi_1(\lambda)+\Pi_2(\lambda))\ln(\Pi_1(\lambda)+\Pi_2(\lambda))\right\}=$$

$$=-2E\left\{\Pi_1(\lambda)\ln\Pi_1(\lambda)\right\}+E\left\{\Pi_1(2\lambda)\ln\Pi_1(2\lambda)\right\}=\mu(2\lambda)-2\mu(\lambda).$$

Отсюда следует, что

$$\mu_T=\frac{\mu(2\lambda)-2\mu(\lambda)}{2\lambda}=\frac{1}{2\lambda}\left(e^{-2\lambda}2\lambda\sum_{k=1}^{+\infty}\frac{(2\lambda)^k\ln(k+1)}{k!}-2e^{-\lambda}\lambda\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}\right)=$$

$$=e^{-\lambda}\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}\left(e^{-\lambda}2^k-1\right).$$

Теперь докажем (21). Справедливо представление дисперсии величины (10):

$$D\left\{g\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)\right\}=D\left\{f\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)\right\}-\gamma^2D\left\{\Pi_1(\lambda)+\Pi_2(\lambda)-2\lambda\right\}=$$
$$=D\left\{f\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)\right\}-2\lambda\gamma^2. \qquad (22)$$

Преобразуем слагаемые (22). Воспользовавшись определением дисперсии и (6), раскрыв скобки и перегруппировав, получим представление первого слагаемого:

$$D\left\{f\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)\right\}=E\left\{f^2\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)\right\}-\left(\mu(2\lambda)-2\mu(\lambda)\right)^2=E\left\{\Pi_1^2(\lambda)\ln^2\Pi_1(\lambda)\right\}+$$
$$+E\left\{\Pi_2^2(\lambda)\ln^2\Pi_2(\lambda)\right\}+E\left\{\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)^2\ln^2\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\right\}+$$
$$+2E\left\{\Pi_1(\lambda)\Pi_2(\lambda)\ln\Pi_1(\lambda)\ln\Pi_2(\lambda)\right\}-$$
$$-2E\left\{\Pi_1(\lambda)\ln\Pi_1(\lambda)\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\ln\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\right\}-$$
$$-2E\left\{\Pi_2(\lambda)\ln\Pi_2(\lambda)\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\ln\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\right\}-$$
$$-\mu^2(2\lambda)-4\mu^2(\lambda)+4\mu(2\lambda)\mu(\lambda)=$$
$$=\left(E\left\{\Pi_1^2(2\lambda)\ln^2\Pi_1(2\lambda)\right\}-\mu^2(2\lambda)\right)+4\left(E\left\{\Pi_1^2(\lambda)\ln^2\Pi_1(\lambda)\right\}-\mu^2(\lambda)\right)-$$
$$-4\left(E\left\{\Pi_1(\lambda)\ln\Pi_1(\lambda)\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\ln\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\right\}-\mu(2\lambda)\mu(\lambda)\right).$$

С учётом (11) и (12) получим

$$D\left\{f\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)\right\}=\sigma^2(2\lambda)+4\sigma^2(\lambda)-4\nu(\lambda). \qquad (23)$$

Для того, чтобы вычислить второе слагаемое, проведём вспомогательные расчёты. Из свойств математического ожидания и (6) следует

$$\operatorname{cov}\left\{f\left(\Pi_1(\lambda),\Pi_2(\lambda)\right),\Pi_1(\lambda)+\Pi_2(\lambda)\right\}=$$
$$=E\left\{\left(f\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)-\mu(2\lambda)+2\mu(\lambda)\right)\left(\Pi_1(\lambda)+\Pi_2(\lambda)-2\lambda\right)\right\}=$$
$$=-E\left\{\Pi_1^2(\lambda)\ln\Pi_1(\lambda)\right\}-E\left\{\Pi_1(\lambda)\Pi_2(\lambda)\ln\Pi_1(\lambda)\right\}+E\left\{2\lambda\Pi_1(\lambda)\ln\Pi_1(\lambda)\right\}-$$
$$-E\left\{\Pi_1(\lambda)\Pi_2(\lambda)\ln\Pi_2(\lambda)\right\}-E\left\{\Pi_2^2(\lambda)\ln\Pi_2(\lambda)\right\}+E\left\{2\lambda\Pi_2(\lambda)\ln\Pi_2(\lambda)\right\}+$$
$$+E\left\{\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)^2\ln\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\right\}-E\left\{2\lambda\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\ln\left(\Pi_1(\lambda)+\Pi_2(\lambda)\right)\right\}.$$

Группируя слагаемые и используя (9), получим

$$\operatorname{cov}\left\{ f\left(\Pi_1(\lambda),\Pi_2(\lambda)\right),\Pi_1(\lambda)+\Pi_2(\lambda)\right\} = -4E\left\{\Pi_1^2(\lambda)\ln\Pi_1(\lambda)\right\} + 4E\left\{\lambda\Pi_1(\lambda)\ln\Pi_1(\lambda)\right\} +$$
$$+E\left\{\Pi_1^2(2\lambda)\ln\Pi_1(2\lambda)\right\} - E\left\{2\lambda\Pi_1(2\lambda)\ln\Pi_1(2\lambda)\right\} =$$
$$= E\left\{\Pi_1(2\lambda)\ln\Pi_1(2\lambda)\left(\Pi_1(2\lambda)-2\lambda\right)\right\} - 4E\left\{\Pi_1(\lambda)\ln\Pi_1(\lambda)\left(\Pi_1(\lambda)-\lambda\right)\right\} = \zeta(2\lambda)-4\zeta(\lambda).$$

Отсюда с учётом (8) следует

$$\gamma = \frac{\zeta(2\lambda)-4\zeta(\lambda)}{2\lambda}. \tag{24}$$

Из (23) и (24) вытекает

$$\sigma_*^2 = \frac{2^{n-1}D\left\{g\left(\Pi_1(\lambda),\Pi_2(\lambda)\right)\right\}}{T^2} = \frac{2^{n-1}}{T^2}\left(\sigma^2(2\lambda)+4\sigma^2(\lambda)-4\nu(\lambda)-\frac{\left(\zeta(2\lambda)-4\zeta(\lambda)\right)^2}{2\lambda}\right). \tag{25}$$

Используя (11) и проводя эквивалентные преобразования, получим

$$\sigma^2(\lambda) = \sum_{k=0}^{+\infty} k^2\ln^2 k P\{\Pi_1(\lambda)=k\} = \sum_{k=0}^{+\infty} k^2\ln^2 k\,\frac{e^{-\lambda}\lambda^k}{k!} - \mu^2(\lambda) =$$
$$= e^{-\lambda}\sum_{k=2}^{+\infty} k\ln^2 k\,\frac{\lambda^k}{(k-1)!} - \mu^2(\lambda) = e^{-\lambda}\lambda\sum_{k=1}^{+\infty}(k+1)\ln^2(k+1)\frac{\lambda^k}{k!} - \mu^2(\lambda); \tag{26}$$

Проводя аналогичные преобразования для (9), имеем

$$\zeta(2\lambda)-4\zeta(\lambda) = \sum_{k=0}^{+\infty} k(k-2\lambda)\ln k\,\frac{e^{-2\lambda}(2\lambda)^k}{k!} - 4\sum_{k=0}^{+\infty} k(k-\lambda)\ln k\,\frac{e^{-\lambda}\lambda^k}{k!} =$$
$$= e^{-2\lambda}\sum_{k=2}^{+\infty}(k-2\lambda)\ln k\,\frac{(2\lambda)^k}{(k-1)!} - 4e^{-\lambda}\sum_{k=0}^{+\infty}(k-\lambda)\ln k\,\frac{\lambda^k}{(k-1)!} = \tag{27}$$
$$= 2\lambda e^{-\lambda}\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}\left(2^k e^{-\lambda}(k+1-2\lambda)-2(k+1-\lambda)\right).$$

Подставив в (25) значения (26), (27), получим

$$\sigma_*^2 = \frac{2^{n-1}}{T^2}\left(2\lambda e^{-2\lambda}\sum_{k=1}^{+\infty}(k+1)\ln^2(k+1)\frac{(2\lambda)^k}{k!} - \mu^2(2\lambda) + 4\lambda e^{-\lambda}\sum_{k=1}^{+\infty}(k+1)\ln^2(k+1)\frac{\lambda^k}{k!} -\right.$$
$$-4\mu^2(\lambda) - 4\lambda e^{-2\lambda}\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}\sum_{j=1}^{k}(j+1)\ln(j+1)C_{k+1}^{j+1} + 4\mu(2\lambda)\mu(\lambda) - \tag{28}$$
$$\left.-2\lambda e^{-2\lambda}\left(\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}\left(2^k e^{-\lambda}(k+1-2\lambda)-2(k+1-\lambda)\right)\right)^2\right).$$

Из (14) следует

$$\mu^2(2\lambda)+4\mu^2(\lambda)-4\mu(2\lambda)\mu(\lambda) = \left(\mu(2\lambda)-2\mu(\lambda)\right)^2 = 4\lambda^2 e^{-2\lambda}\left(\sum_{k=1}^{+\infty}\frac{\lambda^k\ln(k+1)}{k!}\left(2^k e^{-\lambda}-1\right)\right)^2. \tag{29}$$

Также из определения λ справедливо соотношение

$$\frac{2^n\lambda e^{-\lambda}}{T^2} = \frac{e^{-\lambda}}{T}. \tag{30}$$

Подставив в (28) значение (29), с учётом (30) приходим к (21). ■

Отметим, что выражения для величин (23) и (24) приведены в [2] без доказательства и с опечатками.

В [1] описан тест приращения энтропии, который опирается на теорему 1 [2]. Пусть $\varepsilon \in (0,1)$ – заданный уровень значимости. Вычислим для наблюдаемой последовательности длины *T* статистику (5) при некотором *n*. Выносится решение при помощи решающего правила [1]:

$$\begin{cases} H_*, & \text{если } t_- < G(T,n) < t_+; \\ \overline{H}_*, & \text{в противном случае,} \end{cases} \qquad t_\pm = \mu_T \pm \sigma_* \Phi^{-1}\left(\frac{\varepsilon}{2}\right),$$

где $\Phi^{-1}(\cdot)$ – квантиль стандартного нормального закона [3]. При $T \to +\infty$ асимптотический размер этого теста совпадает с ε.

Доказанная нами теорема 2 позволяет применить на практике указанный тест, поскольку нами получены явные формулы математического ожидания и дисперсии (5). Например, этот тест можно использовать как критерий оценки качества криптографических генераторов. Кроме того, теорема 2 позволяет вычислить математическое ожидание оценки энтропии (4).

Теорема 3. При истинной гипотезе H_*

$$E\{\hat{h}(T,n)\} = h_*(T,n) = \sum_{m=1}^{n} e^{-\frac{T}{2^m}} \sum_{k=1}^{\infty} \frac{T^k \ln(k+1)}{2^{mk} k!} (e^{-\frac{T}{2^m}} 2^k - 1). \qquad (31)$$

Доказательство. Из (5) вытекает $h_*(T,n) = G_*(T,n) + h_*(T,n-1)$, $h_*(T,0) = 0$. Просуммировав с учётом (20), получим требуемое. ■

Полученное значение математического ожидания (31) также можно использовать для построения критериев качества криптографических генераторов.

ЛИТЕРАТУРА

1. *Харин, Ю. С.* Криптология / Ю. С. Харин, С. В. Агиевич, Д. В. Васильев, Г. В. Матвеев. Мн : БГУ, 2013. 511 с.

2. *Rukhin, A. L.* Approximate Entropy for Testing Randomnesses / A. L. Rukhin. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.45.6174&rep=rep1&type=pdf.